

CUSTOMER RELATIONSHIP MANAGEMENT IN THE MANUFACTURING INDUSTRY, USING DATA MINING TECHNIQUES

P. RAJENDIRAN, K. R. SEKAR, B. SINIVASAN & R. MANIKANDAN

School of Computing, SASTRA University, Thanjavur, India

ABSTRACT

In the eventual world, production industry tries to get more market share for their products. This is possible only through Customer Relationship Management (CRM). A separate team has been working for the same, to improve the customers need and make the production in the positive direction. Strenuous remedial measure has to be taken to bring up the sales quantity. Various methodologies have to be incorporated, to obtain good strategies for CRM. Data mining is the crisp methodological fragrance, dominated to the core. Naive base, Decision tree, Prism, Ada Boost and Ensemble are the existing methodologies that predominately capture the market more. In our research work, DBSCAN method has to be employed, to obtain and grasp the people to the required segment. Thus, the production company improves their requirement sales, through the people demands.

KEYWORDS: *Online Advertising, Customer Relationship, Productivity & Manufacturing Industry*

Received: Sep 27, 2017; **Accepted:** Oct 16, 2017; **Published:** Oct 30, 2017; **Paper Id:** IJMPERDDEC20176

INTRODUCTION

Any Manufacturing industry needs advertising their product in the market, in a highly pronounced manner. Targeting the customers are their essential requirements in the day to day practice, to improve the sales of the manufacturing products. Big data Analytics is the wonderful tool, to know about the customer behavior and their tendency to consume the products.

Advertising is one of the most widely used forms of marketing. To sell any product, it should be advertised well and more importantly, it should reach the right customer. In the present time, online advertising is the most preferred form of marketing, because of the growth and reach of internet. Most of the people nowadays prefer to buy a product online, as it saves time for them. There have been many online advertising measures that have been adopted from the yester years, but none can claim to be perfect. The reason for none of the methods being perfect is, because of the huge cost involved in it. Advertising is a costly affair these days and it is considered to be a key factor that promotes the sales and growth of a company. So, the ultimate aim of every company is to spend minimum amount on advertising and generate maximum profit. Also, maintaining a cordial relationship with the existing customers is one of the main factors, which many companies fail to do.

LITRATURE SURVEY

The existing system for a hairdressing industry, classifies the customers into three categories. The categories are loyal customer, potential customer and new customer. The main aim in the existing system is, to give the right offers on hairdressing and hair care products for the customers, based on their category in which they lie. The customers are categorized into different categories, based on a method called RFM (Recency, Frequency,

and Monetary). Recency; indicates the date of the latest visit made by the customer. Frequency; indicates the number of times the customer has visited or the number of visits made by the customer, in a period of time. Monetary; indicates the total amount spent by the customer, for the number of visits made in a period of time.

If, all the three factors are high (RFM is high), then such customers are said to be loyal customers. If, all the three factors are medium (RFM is medium), then such customers are said to be Potential customers. If, all the three factors are low (RFM is low), then such customers are said to be new customers.

SOM and K-Means are the clustering methods used to classify the customers into different groups. SOM is an unsupervised learning mechanism in Artificial Neural Network, which is used for problem solving and market screening. It helps in detecting the strong and weak features, in a large dataset. K-Means is used for clustering and placing the points in the closest region, and then recalculates centroid distance based on the points assigned.

EXISTING METHODS AND DRAWBACKS

The existing method in the manufacturing industry, classifies the customers into only one category, wherein all the customers get the same offers and discounts. In this case, if a customer did not like the offer, the probability that he purchases becomes less. The customer might visit some other shop for some exciting offers. This might result in losing the existing customer, which is one of the main drawbacks in this method. Also, a customer who is interested in a particular product will get unwanted advertisements, which the customer is not interested in.

This results in a drawback, wherein the cost of advertising becomes higher and where the probability that the customer will purchase the product is displayed in the advertisement. All these factors affect the Customer relationship and decreases [12] the profit of the company to a great extent, which is another drawback.

PROPOSED METHOD AND ITS ADVANTAGE

BEHAVIORAL TARGETING

The proposed method involves using Behavioral Targeting [4], which is a technique used by online advertisers and merchants, to efficiently advertise and market their products through internet. The individual searching behaviour is extracted and stored in the database. Using Behavioral targeting, we can create a complete profile for the customers, along with their purchasing history. The customer profile is stored in the database. RFM [13] (Recency, Frequency, Monetary) is a technique, which is used along with Behavioral Targeting, to classify the customers into three different types. The three different types of customers are [1]

- Loyal customers
- Potential customers
- New customers

After extracting the user's interest, the user will be asked to enter the Id. Once the Id is entered, it is checked in the database to know whether the Id exists and to know what type of customer he is. If the id exists and the type of customer is known, the advertisements are displayed according to their type.

One of the main advantages with this technique is, it provides different set of advertisements, based on the type of customer. Every time the user searches for a product, the keyword he searches is stored in the database as his interest.

This interest column keeps on updating, every time the user logs.

After the interest reaches a particular level, it selects the product that has been searched the most and the next time when the user logs in, the search string is not taken for the cloth type, rather the given id is matched with the database and the advertisements are displayed for the product, which has been searched the most number of times by that user. One of the main advantages of this method is, it helps in reducing the amount spent for the advertisements.

CREDIT PAYMENT SYSTEM

The Credit Payment System is a method which avoids internet banking. It is a method by which, the user will be asked to store some amount of money in credit. Every customer will be given a Smart card, during their first purchase. It is a simple one step process, by which the user can buy the product in a single click of a button. The details of credit for every user are stored in the database. When the user purchases a product from the list of advertisements displayed in his page, the amount for the product chosen by the user will be taken from his credit. If, the credit amount is less than the amount needed for the product, then the payment process is discarded. The payment process occurs successfully, by matching the id entered by the user to his credit.

The users who visit the page for the first time cannot make use of this system and hence, has to use internet banking only, for the first time. The advantages of this method are that, it is very simple and more secure when compared to internet banking. Adding to that, the transaction is also faster and there is no chance for any issues like session expiry or server problem.

MAPPING OF FREQUENCY AND MONETARY

Mapping is a method to determine the number of users in various categories. It is done taking frequency along the x-axis and monetary along the y-axis. [14] Based on the frequency and monetary of every user, the points are plotted in the graph.

The details of frequency and monetary for all the users is taken from the database. The frequency range is different for different categories of customers. They are

- The frequency range for a Loyal Customer is >7 .
- The frequency range for a Potential Customer is >4 .
- The frequency range for a New Customer is between 1 to 3.

Once, the points are plotted, we get a graphical representation of how many users exist in various segments. Based on the mapping, the customer who does not fall into general category of purchasing is classified as

- Small Scale Loyal Customer (if the Monetary range is less than 60,000)
- Small Scale Potential Customer (if the Monetary range is less than 40,000)
- Small Scale New Customer (if the Monetary range is less than 15,000)

The main advantage of mapping is, it helps in determining the growth in sales and revenue of a company in a period of time, as well as finding the customers who frequently purchase and those customers who rarely purchase.

DBSCAN ALGORITHM

DBSCAN (Density Based Spatial Clustering of Application with Noise) is a [6] spatial clustering algorithm, in which the cluster is defined based on density reachability. DBSCAN

Algorithm takes only two parameters as input. They are [15]

- Radius (r)
- Minimum Points needed for forming clusters (minimum pts)

The coordinate points obtained from mapping are taken as input, for which the DBSCAN [6] algorithm is to be performed. At first, the distance of one point with every other point is calculated, using the usual distance formula. The DBSCAN algorithm classifies the points into three different categories. They are [7]

- Core points
- Boundary points
- Outliers

If, the distance of a point is less than the radius, then the number of neighbours for that point is determined, by comparing the distance with all the other points.

A point is said to be a Core point, if the number of neighbours is greater than or equal to the minimum pts.

A point is said to be a Boundary point, if the number of neighbours is less than to the minimum pts. A point is said to be an Outlier, if the number of neighbours is less than to the minimum pts and if, the point is not a boundary point.

All the Core points lie inside the cluster, whereas all the [8] Boundary points lies exactly on the cluster and all the outliers fall outside the cluster.

MAILING THE SAMLL SCALE CUSTOMERS

All the Small Scale Customers are sent exciting gifts and special offers through email [3]. This technique is followed, mainly to attract the customers and to increase their purchasing habit. The mail id of the customers is taken from the database. Mail is sent using a simple SMTP protocol.

The main aim of mailing the customers is, to bring all these Small Scale Customers [9] into the general category. One of the main advantages of mailing is, it helps in maintaining a healthy relationship with the existing customers. The customers can also mail in case of any query, to the mail specified by the company [10].

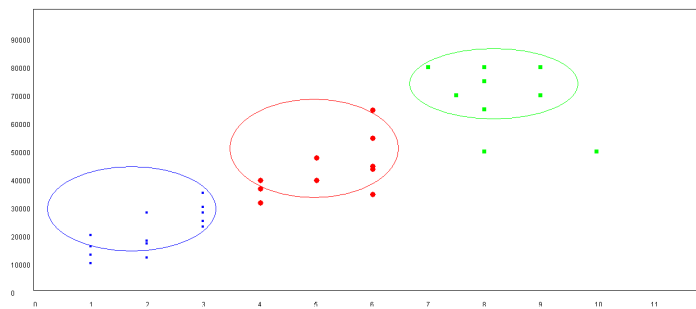
RESULTS AND ANALYSIS

Customer Profile in Database

```
MySQL5.5 Command Line Client
mysql> use backend;
Database changed
mysql> select * from cust_details;
```

id	name	credit	interests	recency
frequency	monetary	age	sex	mailid
				type
114015001	raj	9000	tshirt	2014-03-25
8	80000	23	m	rajraj@gmail.com
114015002	siva	29000	jeans,shirt,tshirt,jeans,jeans	2014-03-18
10	50000	22	m	rajendiranap@gmail.com
114015003	ravi	12000	shirt,shirt,jeans,jeans,shirt	2014-03-19
8	70000	26	m	pavil2@gmail.com
114015004	srikan	9000	shirt,tshirt	2014-03-19
8	50000	20	f	srikan92rockin@gmail.com
114015005	pooja	14500	tshirts	2014-01-08
7	80000	22	f	pooja34@gmail.com
114015007	priya	9201	shirts,shirt	2014-03-25
7	80000	22	f	priya1234@gmail.com
114015008	ashok	16000	jeans	2013-12-26
7	85000	22	m	ashokultimate@gmail.com
114015009	vikran	12500	jeans	2014-01-20
9	70000	22	m	vikran8877@gmail.com
114015039	harry	15700	jeans	2013-11-23
9	70000	21	m	harryboy@gmail.com
114015045	srikan	13800	jeans	2014-01-01
8	75000	45	m	srikan92rockin@gmail.com
114015046	harsh	14000	shirt	2014-02-02
9	80000	25	m	harshmail@gmail.com
114015047	srishavin	15770	tshirt,jeans	2014-03-25
10	70000	18	m	srishavin22@gmail.com
114015048	visakan	10000	jeans	2014-03-12
8	65000	28	m	visak44@gmail.com
114015049	hansika	10500	jeans,jeans	2014-03-25
11	50000	28	f	hanshika17@gmail.com
114015050	anilrudh	8550	jeans,tshirt,jeans	2014-03-25
9	50000	28	m	aninol7@gmail.com
200004001	vijay	3500	shirts,shirt	2014-03-25
6	65000	35	m	vij333@gmail.com
200004002	narayan	9600	tshirt,jeans	2014-03-25
5	40000	43	m	naray956@gmail.com

MAPPING



DBSCAN CLUSTERING

```

neighbours presents are 6
Core point : (900,200)
neighbours presents are 7
Core point : (850,250)
neighbours presents are 2
neighbours presents are 6
Core point : (900,200)
neighbours presents are 6
Core point : (900,213)
neighbours presents are 4
neighbours presents are 4
neighbours presents are 7
Core point : (900,275)
neighbours presents are 3
neighbours presents are 3
neighbours presents are 4
neighbours presents are 4
neighbours presents are 7
Core point : (900,275)
neighbours presents are 3
neighbours presents are 3
Core point : (900,200)
Core point : (850,250)
Boundary point: (900,200)
Core point : (900,200)
Core point : (900,225)
Boundary point: (1000,200)
Core point : (900,275)
Boundary point: (1000,200)
Core point : (900,275)
Boundary point: (900,350)
Boundary point: (900,300)
Boundary point: (1000,200)
Outlier : (1000,250)
Outlier : (1100,350)
Outlier : (1100,350)
Boundary point: (1000,200)
Outlier : (1000,250)
Outlier : (1100,350)

```

CONCLUSIONS

There are various ways in which, online advertising can be compromised. Advertising a product for a user, who does not need that product, ends up in money being spent uselessly. To avoid this, behavioral targeting has been used. With the fast improvement in the manufacturing industry, Customers are classified into different categories, based on RFM technique [11]. When it comes to Behavioral targeting, the right advertisements are displayed for the right customers, based on their purchasing capacity. In addition to this, mapping is performed to differentiate the normal and small scale customers. Based on the mapped points, clustering is performed by DBSCAN algorithm, which is an efficient method for use with database. By using the above mentioned methods, the manufacturing industry can significantly improve its profit and maintain a cordial relationship with the customers, at the same time winning new customers. These techniques can be further extended to other commercial fields of business, which involves lots of customers.

REFERENCES

1. Jo-Ting Wei Ming-Chun Lee Hsuan-Kai Chen Hsin-Hung Wu On Customer relationship management in the hairdressing industry: An application of data mining techniques, *Expert Systems with Applications*, Volume 40, 7513–7518, 2013.
2. Brown U. J Beale R. L On “Services Marketing: The mediating role of customers in the hair care industry”, *Academy of Marketing Studies Journal*, 12(1), 57-70, 2008.
3. Swift R. S On “Accelerating Customer relationship – Using CRM and relationship technologies, *International Journal of Organizational Innovation*, 2(4), 104-125, 2001.
4. Li Jun Zhang Peng On “Mining Explainable User Interests from Scalable User Behaviour Data”, *Elsevier*, 789-796, 2013.
5. ZHOU Shui Geng, ZHOU Ao Ying, and CAO Jing On “A Data Partitioning Based DBSCAN Algorithm”, *Journal on Computer Research and Development*, 2004.
6. Neaga, E. I., and Harding, J. A., 2002, “A Review of Data Mining Techniques and Software Systems to Improve Business Performance in Extended Manufacturing Enterprises,” *Int. J. Adv. Manuf. Syst. (IJMAS)*, Spec Issue Decis. Eng., 5 (1), pp. 3–19.
7. Neaga, E. I., and Harding, J. A., 2001, “Data Mining Techniques for Supporting Manufacturing Enterprise Design,” *International Conference on Industrial and Production Management*, Quebec City, Canada, pp. 232–241.
8. SPSS, 1999, “CRISP-DM 1.0 Step-By-Step Data Mining Guide”.
9. Shahbaz, M., 2005, “Product and Manufacturing Process Improvement Using Data Mining,” Ph. D. thesis, Wolfson School of Mechanical and Manufacturing Engineering, Loughborough University, Loughborough, Leicestershire, 194 pp.
10. Neaga, E. I., and Harding, J. A., 2005, “An Enterprise Modelling and Integration Framework Based on Knowledge Discovery and Data Mining,” *Int. J. Prod. Res.*, 43 (6), pp. 1089–1108.
11. Neaga, E. I., 2003, “Framework for Distributed Knowledge Discovery Systems Embedded in Extended Enterprise,” Ph. D. thesis, Wolfson School of Mechanical and Manufacturing Engineering, Loughborough University, Loughborough, UK, 192 pp.
12. Sim, S. K., and Chan, Y. W., 1992, “A Knowledge-Based Expert System for Rolling-Element Bearing Selection in Mechanical Engineering Design,” *Artif. Intell. Eng.*, 6 (3), pp. 125–135.
13. Kusiak, A., Kernstine, K. H., Kern, J. A., McLaughlin, K. A., and Tseng, T. L., 2000, “Data Mining: Medical and Engineering Case Studies,” *Industrial Engineering Research Conference*, Cleveland, OH, pp. 1–7.

14. Ishino, Y., and Jin, Y., 2001, "Data Mining and Knowledge Acquisition in Engineering Design," *"Data Mining for Design and Manufacturing: Methods and Applications"*, D. Braha, ed., Kluwer Academic, Dordrecht, pp. 145–160.
15. Romanowski, C. J., and Nagi, R., 2001, "A Data Mining for Knowledge Acquisition in Engineering Design," *"Data Mining for Design and Manufacture: Methods and Applications"*, D. Braha, ed., Kluwer Academic, Dordrecht, pp. 161–178.

